IBM Cloud Pak for Data System

Quickly scale and deploy a complete private cloud for your enterprise data and AI architecture

For years, companies have been accumulating data at an impressive pace. Some of this data is being duplicated across various clouds and repositories at astounding levels. This duplication leads to data silos, high costs, delayed projects and increased security risks.

Data users assigned to an analytics initiative spend as much as 60 percent of their overall time trying to locate the data they need. Organizations have existing processes to store data in multiple repositories, clouds and backend systems and have invested resources to move data to a central location where it can be managed, controlled and made accessible. However, most centralized big data projects still fail to deliver on the promise of easy and controlled data access. Turning that failure to success would allow data scientists to access data across their organization wherever it resides. This greatly reduces the amount of time spent searching for data.

As such, IBM developed Cloud Pak for Data System, a hyper-converged system that combines storage, computing and networking software into a single system. Built on a governed Data and AI platform, Cloud Pak for Data System simplifies and unifies how you manage, govern and analyze data. It allows you to provision and deploy data services flexibly and rapidly, tailored to your needs. You can automatically discover and add hardware nodes and add software from an app-like store to expand your system.

Cloud Pak for Data System installation is designed to be as simple as using the software itself.

Highlights
• Rely on a data platform with built-in governance
• Simplify how you manage, govern and analyze data
• Adapt to your needs using highly available services
• Gain instant preassembled software provisioning capabilities with hardware
Hyper-converged technology makes it possible for users to scale and evolve their infrastructures simply and economically as application loads change. Cloud Pak for Data System includes a flexible microservices software architecture, field-programmable gate arrays (FPGAs) and FPGA-enabled network interface cards (NIC) for advanced AI acceleration. Start scaling for your needs and deploying faster with Cloud Pak for Data System.

**Accelerate time to value with IBM Cloud Pak for Data System**

Customers are aware of the value and speed cloud provides to their applications and services, and want the same efficiency in their own data centers. Cloud Pak for Data System offers accelerated time to value, allowing you to stand up an entire private cloud system for your data and AI architecture in under four hours.

Hyper-convergence enables a software-defined data center to provide cost-effective agility, scalability and security. Customers can quickly develop applications and analyze their data in a security-rich environment within their own firewalls and expand as they need rather than resort to expensive capital outlays. Most importantly, developers don’t need to change how they deploy applications, whether they’re working in the cloud or their own data center.

When your needs outgrow your existing installation, Cloud Pak for Data System simplifies expansion. Buy small increments of computing or storage capacity as needed, rather than outlaying vast amounts of capital. Based on application demand and delivery, budgeting can be completed in quarterly expansions on an ad hoc basis. When the hardware arrives, it’s pre-loaded and configured with software and licenses. Once installed in the data center, the existing cluster will automatically discover new hardware and licenses, inherit all user configurations, and be ready to use in a few hours. The platform evolves based on your needs and scalability, and it’s all managed and highly available to help ensure maximum performance and productivity.

**Bring Data & AI Workloads to Red Hat OpenShift Environment**

With your Docker registry, your tools will always run on Cloud Pak Data. Cloud Pak for Data System is natively built with Red Hat OpenShift Container Platform. Containers built to run on any OpenShift environment migrate seamlessly to Cloud Pak for Data – and on any OpenShift cloud or in the System.

Cloud Pak for Data supports three types of registries: Red Hat, its own, and those in your organization. With OpenShift, you can easily write and deploy applications knowing that they’ll run on a platform optimized for Red Hat OpenShift. When choosing to deploy a private cloud on-premises, Cloud Pak for Data System provides optimized hardware to increase the container performance of the Red Hat cluster while speeding the time to value of data workloads.
IBM Performance Server for PostgreSQL

As part of Cloud Pak for Data System, IBM Performance Server for PostgreSQL is 100% compatible with IBM PureData System for Analytics including FPGA query acceleration. This allows customers to combine their existing PureData System for Analytics in-database and machine learning applications with Cloud Pak for Data’s powerful AI capabilities to enrich their business with insight and AI.

For years, companies have accumulated and analyzed data in PureData System for Analytics and now they can benefit from this new hyper-converged platform for speed and simplicity. When provisioned in an existing Cloud Pak for Data System, Performance Server extends the capabilities of PureData and includes all the additional benefits of Cloud Pak for Data to consolidate data and reduce data duplication. PureData users familiar with Fluid Query will find it complements Cloud Pak for Data’s capabilities, allowing access to other data sources to eliminate data silos, high costs, delayed projects and high security risks.
**Why IBM?**

IBM's open information architecture for AI is built upon Cloud Pak for Data, which supports:

- Cloud-native architecture, including containerized workloads, microservices, and multicloud provisioning.

- Cloud open source platforms/applications such as Docker and Kubernetes.

- Open source frameworks and databases such as Apache Hadoop, MongoDB.

- AI open source programming languages, deep learning frameworks, and open source interfaces.

**Next steps**

Get started by learning about Cloud Pak for Data at [ibm.biz/CloudPak4Data](http://ibm.biz/CloudPak4Data), or by trying the [7-day free trial](http://ibm.biz/CloudPak4Data).

You can also [schedule a free consultation](http://ibm.biz/CloudPak4Data) for more information tailored to your business needs.

---

**For more information**

To learn more about IBM Cloud for Data Pak System, visit [ibm.biz/CloudPak4DataSystem](http://ibm.biz/CloudPak4DataSystem).